

日本国特許庁
JAPAN PATENT OFFICE

1-18-02

Priority
Papers



別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office

出願年月日

Date of Application: 2000年 9月 5日

出願番号

Application Number: 特願 2000-267954

出願人

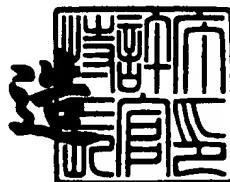
Applicant(s): バイオニア株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 8月 3日

特許庁長官
Commissioner,
Japan Patent Office

及川耕



出証番号 出証特 2001-3070021

【書類名】 特許願
【整理番号】 54P0632
【提出日】 平成12年 9月 5日
【あて先】 特許庁長官殿
【国際特許分類】 G10L 3/00
G08G 1/969

【発明者】

【住所又は居所】 埼玉県川越市山田字西町25番地1 パイオニア株式会
社川越工場内

【氏名】 斎藤 宏

【特許出願人】

【識別番号】 000005016
【氏名又は名称】 パイオニア株式会社

【代理人】

【識別番号】 100063565

【弁理士】

【氏名又は名称】 小橋 信淳

【手数料の表示】

【予納台帳番号】 011659
【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声認識装置ならびにその方法

【特許請求の範囲】

【請求項1】 階層的に互いに関連づけられた複数の音声認識辞書と、前記音声認識辞書のうち所望の辞書を待ち受けワード一覧として抽出する抽出手段と、

前記音声認識辞書のうち、所望の辞書の選択操作を行う選択手段と、

前記選択手段で選択された辞書をあらかじめ設定されている階層より上位階層の待ち受けワード一覧とし、前記抽出手段で抽出された通常の辞書と併せて記憶する記憶手段と、

入力された音声と前記記憶手段に記憶された待ち受けワード一覧とを比較することによって入力された音声を認識する認識手段とを備えたことを特徴とする音声認識装置。

【請求項2】 前記音声認識辞書として、施設の種別名を記憶した種別辞書と、施設の種別に属する施設の施設名を種別毎に記憶した施設辞書とを備えることを特徴とする請求項1に記載の音声認識装置。

【請求項3】 前記音声認識辞書として、地域名を記憶した地域辞書と、いずれかの地域に存在する施設の施設名を地域毎に記憶した施設辞書とを備えることを特徴とする請求項1または2に記載の音声認識装置。

【請求項4】 前記選択手段は、前記施設辞書を所望の辞書として選択することを特徴とする請求項2または3に記載の音声認識装置。

【請求項5】 前記抽出手段は、認識された音声の下位階層の辞書を待ち受けワードとして抽出するとともに、前記選択手段で選択された辞書に属する辞書で、且つ認識された音声の下位階層の辞書を待ち受けワードとして抽出することを特徴とする請求項4に記載の音声認識装置。

【請求項6】 階層的に互いに関連付けられた音声認識辞書を複数持つ音声認識装置に用いられ、

上位階層において絞り込み条件の辞書と共に、あらかじめユーザにより設定される少なくとも一つの絞込み条件によって区分された辞書を認識の対象として入

力された音声の認識処理を行うことを特徴とする音声認識方法。

【請求項7】 前記あらかじめユーザにより設定される少なくとも一つの絞込み条件によって区分された辞書は、使用頻度の高い辞書であることを特徴とする請求項6に記載の音声認識方法。

【請求項8】 階層的に互いに関連付けられた複数の音声認識辞書と、前記音声認識辞書のうち所望の辞書を待ち受けワード一覧として抽出する抽出手段と、

前記抽出手段で抽出された辞書の待ち受けワード一覧を記憶する記憶手段と、入力された音声と前記記憶手段に記憶された待ち受けワード一覧とを比較することにより入力された音声を認識する認識手段とを備え、

前記認識手段によって音声が認識された際に、前記抽出手段が認識された音声の下位階層の辞書の待ち受けワードとして抽出し前記記憶手段が記憶するとともに、認識される際に前記記憶手段に記憶されていた待ち受けワードのうち認識された音声に関連する待ち受けワードを引き続き比較対象として記憶しておくことを特徴とする音声認識装置。

【請求項9】 階層的に互いに関連付けられた複数の音声認識辞書から、所望の辞書を待ち受けワード一覧として抽出し、抽出された辞書の待ち受けワード一覧を記憶し、入力された音声と記憶された待ち受けワード一覧とを比較することにより入力された音声を認識する音声認識方法であって、

音声が認識された際に、認識された音声の下位階層の辞書を待ち受けワードとして抽出し記憶するとともに、認識される際に記憶されていた待ち受けワードのうち認識された音声に関連する待ち受けワードを引き続き比較対象として記憶しておくことを特徴とする音声認識方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、操作性ならびに応答性の向上をはかった音声認識装置ならびにその方法に関する。

【0002】

【従来の技術】

従来、音声認識装置を用いて施設名称等を検索する場合、認識率の確保や使用可能なメモリサイズ等の制約から、図13に絞込みの手順を示すように、ジャンルや地名によって待受け単語を絞り込んだ上で最終的に名称を発話している。

ここでいう音声認識とは、例えば、カーナビゲーションにおいて、ユーザがマイクから入力した音声を認識し、認識した音声を用いて操作処理を行なう音声操作のための音声認識であり、特に、膨大な施設候補のうちから所望の施設を選択する操作を音声で行う音声認識をいう。初期の段階では、システム側では、カーナビゲーションを操作するための制御コマンド辞書が設定されており、ユーザが「目的地設定」とコマンドを発話することでシステムに目的地までの経路設定を行う意思を伝える。

【0003】

システム側では目的地となる具体的な場所を検索する必要があるが、施設数が膨大であるため、一度の音声認識で特定することは不可能である。そこで、検索対象となる施設数を減少させるため、ジャンル名での絞込みを行う。はじめにジャンルでの絞込みを行うため、認識辞書をジャンル名辞書に入れ替えた後、①「ジャンル名をどうぞ」と発話し、ユーザに対してジャンル名の発話を促す。これに対してユーザが②「教育施設」と発話すると音声認識装置が発話を認識する。教育施設ではさらに詳細なサブジャンルを指定するように、システム側で決められており認識辞書をサブジャンル名辞書に入れ替えた後、③「次のジャンル名をどうぞ」と発話し、ユーザに対してサブジャンル名の発話を促す。これに対してユーザが④「高校」と発話すると音声認識装置が発話を認識する。

【0004】

サブジャンルが確定すると、次に地域での絞込みを行うため、認識辞書を都道府県名辞書に入れ替えた後、⑤「都道府県名をどうぞ」と発話し、ユーザに対して都道府県名単位での地域絞込みを促す。これに対してユーザが、⑥「東京都」と発話すると音声認識装置が発話を「東京都」と認識する。サブジャンルが「高校」でかつ都道府県名が「東京都」だった場合、市区町村名を指定するように、あらかじめシステム側で決められており、認識辞書を市区町村名辞書に入れ替えた後、⑦「

市区町村名をどうぞ」と発話し、ユーザに対して市区町村名の発話を促す。これに対してユーザが⑧「渋谷区」と発話すると音声認識装置が発話を認識する。ここまで指定すれば施設数が十分絞り込まれているため、施設名称の検索を開始する。

【0005】

システム側は認識辞書を東京都渋谷区の高校辞書に入れ替えた後、⑨「名称をどうぞ」と発話し、施設名称の発話を促す。ここで、ユーザが「〇〇学園」と発話すると音声認識装置が発話を認識し〇〇学園を目的地に設定する。

【0006】

【発明が解決しようとする課題】

このように音声認識辞書の階層構造を順に辿り、全ての絞り込み条件を決定するという煩わしい手順を実行することが必要となる。上述した煩わしい手順の実行を回避するために、最終的に検索したい施設名称等を全て最上位階層で待ち受けるという手法が存在する。

【0007】

しかしながら、この場合、膨大な容量を持つメモリが必要となり、また、認識率が低下してレスポンス性能が悪いといった問題もあった。例えば、あるユーザはゴルフをしないため、ゴルフ場を検索することはないが、ユーザにとって興味のないジャンル（この場合、ゴルフ場）も含めた全ての施設名称を待ち受けている場合には、ゴルフ場の名称として誤認識されることがありえる。これはユーザにとって大きなストレスとなる。

【0008】

本発明は上記事情に鑑みてなされたものであり、最上位階層において絞り込み条件の辞書の他に、あらかじめユーザにより設定される少なくとも一つの絞込み条件によって区分された辞書を認識の対象として認識処理を行うことにより操作性の改善をはかると共に、レスポンスの向上をはかった音声認識装置ならびにその方法を提供することを目的とする。

【0009】

また、ユーザが頻繁に使用する絞り込み条件（ジャンルや地域名等）をあらか

じめ設定しておくことにより、この絞り込み条件に合致する施設名称に関しては、階層構造を順に辿り絞り込み条件を決定するという煩わしい処理を行うことなく、1回の発声で目的とする施設名称を検索することができ、更に、同時に絞り込み条件辞書も認識対象とするため、あらかじめ設定してある絞り込み条件に合致しない施設名称を検索する必要が生じた場合でも、従来の階層構造を順に辿り絞り込み条件を決定していく手順により検索可能とする音声認識装置ならびにその方法を提供することも目的とする。

【0010】

【課題を解決するための手段】

上述した課題を解決するために、請求項1に記載の発明は、階層的に互いに関連づけられた複数の音声認識辞書と、音声認識辞書のうち所望の辞書を待ち受けワード一覧として抽出する抽出手段と、音声認識辞書のうち、所望の辞書の選択操作を行う選択手段と、選択手段で選択された辞書をあらかじめ設定されている階層より上位階層の待ち受けワード一覧とし、抽出手段で抽出された通常の辞書と併せて記憶する記憶手段と、入力された音声と記憶手段に記憶された待ち受けワード一覧とを比較することによって入力された音声を認識する認識手段とを備えることとした。

【0011】

また、請求項2に記載の発明は、請求項1に記載の音声認識装置において、音声認識辞書として、施設の種別名を記憶した種別辞書と、施設の種別に属する施設の施設名を種別毎に記憶した施設辞書とを備えることとした。更に、請求項3に記載の発明は、請求項1または2に記載の音声認識装置において、音声認識辞書として、地域名を記憶した地域辞書と、いずれかの地域に存在する施設の施設名を地域毎に記憶した施設辞書とを備えることとした。

【0012】

また、請求項4に記載の発明は、請求項2または3に記載の音声認識装置において、選択手段は、前記施設辞書を所望の辞書として選択することとした。更に、請求項5に記載の発明は、請求項4に記載の音声認識装置において、抽出手段は、認識された音声の下位階層の辞書を待ち受けワードとして抽出するとともに

、選択手段で選択された辞書に属する辞書で、且つ認識された音声の下位階層の辞書を待ち受けワードとして抽出することとした。

【0013】

上記構成により、階層構造を持つ音声認識辞書の検索時、最上位階層における絞り込み条件辞書と共に、あらかじめユーザが設定した少なくとも一つの絞り込み条件によって区分された辞書も認識対象として認識処理が行なわれる。すなわち、ジャンルや地域名等ユーザが頻繁に使用する絞り込み条件をあらかじめ設定しておけば、この絞り込み条件に合致する施設名称に関しては、階層構造を順に辿り絞り込み条件を決定するという煩わしい処理を行うことなく、1回の発声で目的とする施設名称を検索可能な音声認識装置を提供できる。また、同時に絞り込み条件辞書も認識対象としているので、あらかじめ設定絞り込み条件に合致しない施設名称を検索する必要が生じた場合でも、従来の階層構造を順に辿り絞り込み条件を決定していく手順により検索可能とする音声認識装置も提供できる。

【0014】

請求項6に記載の音声認識方法は、階層的に互いに関連づけられた音声認識辞書を複数持つ音声認識装置に用いられ、最上位階層において絞り込み条件の辞書と共に、あらかじめユーザにより設定される少なくとも一つの絞込み条件によって区分された辞書を認識の対象として入力された音声の認識処理を行うこととした。また、請求項7に記載の発明は、請求項6に記載の音声認識方法において、あらかじめユーザにより設定される少なくとも一つの絞込み条件によって区分された辞書は、使用頻度の高い辞書であることとした。

【0015】

このことにより、最上位階層において絞り込み条件の辞書と共に、あらかじめユーザにより設定される少なくとも一つの絞込み条件によって区分された辞書を認識の対象として認識処理を行うことによって操作性の改善がはかれる。ジャンルや地域名等ユーザが頻繁に使用する絞り込み条件をあらかじめ設定しておくことにより、この絞り込み条件に合致する施設名称に関しては、階層構造を順に辿り絞り込み条件を決定するという煩わしい処理を行うことなく、1回の発声で目的とする施設名称を検索でき、操作性ならびに応答性の向上がはかれる。

【0016】

請求項8に記載の発明は、階層的に互いに関連付けられた複数の音声認識辞書と、前記音声認識辞書のうち所望の辞書輪待ち受けワード一覧として抽出する抽出手段と、前記抽出手段で抽出された辞書の待ち受けワード一覧を記憶する記憶手段と、入力された音声と前記記憶手段に記憶された待ち受けワード一覧とを比較することにより入力された音声を認識する認識手段とを備え、前記認識手段によって音声が認識された際に、前記抽出手段が認識された音声の下位階層の辞書の待ち受けワードとして抽出し前記記憶手段が記憶するとともに、認識される際に前記記憶手段に記憶されていた待ち受けワードのうち認識された音声に関連する待ち受けワードを引き続き比較対象として記憶しておくこととした。

【0017】

請求項9に記載の発明は、階層的に互いに関連付けられた複数の音声認識辞書から、所望の辞書を待ち受けワード一覧として抽出し、抽出された辞書の待ち受けワード一覧を記憶し、入力された音声と記憶された待ち受けワード一覧とを比較することにより入力された音声を認識する音声認識方法であって、音声が認識された際に、認識された音声の下位階層の辞書を待ち受けワードとして抽出し記憶するとともに、認識される際に記憶されていた待ち受けワードのうち認識された音声に関連する待ち受けワードを引き続き比較対象として記憶しておくこととした。

【0018】

【発明の実施の形態】

図1は、本発明の音声認識装置の一実施形態を示すブロック図である。

図1において、マイク100は、ユーザの発生を集音し、電気信号に変換し、特微量計算部101へ供給する。特微量計算部101は、PCM (Pulse n Modulation) データを音声認識に適した特微量に変換し、認識部102へ供給する。認識部102は、特微量に変換された入力音声とRAM103にロードされた認識辞書中の待ち受けワードとの類似度を計算し、類似度の高いものからN個の待ち受けワード、及びそれぞれの類似度（スコア）を結果として制御部107に出力する。

【0019】

認識辞書記憶部105は、複数の音声認識用の辞書を記憶しているものである。辞書の種類として、複数の絞り込み条件毎に設けられた絞り込み条件辞書と、絞り込み条件の組み合わせによって分類される最終的な地名名称、例えば、具体的な施設名称等を記憶した施設名辞書とがある。更に、絞り込み条件辞書としては、地点検索を行うための都道府県名等広い地域を示す地域名称を記憶した広地域辞書、それぞれの都道府県毎に設けられ、各都道府県に属する市町村名等狭い地域を示す地域名称を記憶した狭地域辞書、施設の種類等検索地点の大分類ジャンル名称を記憶したジャンル辞書、及び、それぞれの大分類ジャンル毎に設けられ、各大分類ジャンルに属するサブジャンル名称を記憶したサブジャンル辞書等がある。

【0020】

認識辞書選択部104は、制御部107からの指令により、認識辞書記憶部105から所望の辞書を選択し、待ち受けワードとしてRAM103にロードする。初期設定部108は、絞り込み条件の組み合わせに属する施設名称辞書のうち所望の辞書を最上位階層の辞書として設定するために、ユーザが選択操作を行うリモコンキーもしくは音声操作手段とからなる。この初期設定部108にて、設定された施設名称辞書がユーザによる初期設定辞書となる。なお、設定方法については後述する。初期設定記憶部106は、初期設定部108によってユーザが初期設定として設定した絞り込み条件もしくはユーザがいずれの施設名称辞書を初期設定辞書として設定したかを記憶しておくものである。

【0021】

音声合成部109は、ガイダンスマッセージやエコーバック用の合成音声を作成し、スピーカ112に出力する。検索部111は、図示せぬ地図データ等のデータベースを備えており、音声認識によって最終的に検索された施設の、所在地図や住所や電話番号やサービス内容等の詳細情報データベースから検索する。結果表示部110は、音声操作を行う際の認識結果、待ち受けワード、ガイダンスマッセージ、エコーバック等とともに、検索部111によって検索された詳細情報表示するディスプレイである。

【0022】

制御部107は、上述した各構成から出力される出力結果に応じて各構成を制御するものである。即ち、制御部107は、音声認識による施設検索がなされる際に、まず、ジャンル辞書を認識辞書選択部104が認識辞書記憶部105から取り出し、RAM103に待ち受けワードとして設定するように制御する。このとき、あらかじめユーザにより設定されている絞り込み条件もしくは施設名称辞書を、初期設定記憶部106を参照することにより認識し、該当する絞り込み条件もしくは施設名称辞書を、同様に認識辞書選択部104が認識辞書記憶部105から抽出し、RAM103に待ち受けワードとして設定するように制御する。

【0023】

ここで、音声合成部109に、例えば、「ジャンル名をどうぞ」とのガイダンスマッセージを作成させ、スピーカ112から出力させる。

【0024】

RAM103に待ち受けワードとして記憶されているジャンルの辞書内の待ち受けワードを音声入力すると、入力された音声が示すジャンルに属するサブジャンルの辞書が認識辞書記憶部105の中から読み出され、RAM103にロードされ、次の待ち受けワードとなる。そして、RAM103に待ち受けワードとして記憶されているサブジャンルの辞書内の待ち受けワードを音声入力すると、入力された音声が示すサブジャンルを記憶しておくとともに、そのサブジャンルに関連付けられた広地域の辞書が、認識辞書記憶部105の中から読み出され、RAM103にロードされ、次の待ち受けワードになる。

【0025】

そして、RAM103に待ち受けワードとして記憶されている広地域の辞書内の待ち受けワードを音声入力すると、入力された広地域に属する狭地域の辞書が認識辞書記憶部105の中から読み出され、RAM103にロードされ、次の待ち受けワードとなる。そして、RAM103に待ち受けワードとして記憶されている狭地域の辞書内の待ち受けワードを音声入力すると、入力された音声が示す狭地域を記憶しておくとともに、その狭地域に関連付けられた具体的な1の地点を示す辞書が認識辞書記憶部105の中から読み出され、RAM103にロード

され、次の待ち受けワードになる。このように、順次待ち受けワードを構成する辞書が変更して階層的に用いられるように、階層構造をなして認識辞書記憶部105に記憶されている。すなわち、後述する図2乃至図5に階層辞書ツリーとして示すように、ジャンルの辞書の下にサブジャンルの辞書があり、広地域の辞書の下に狭地域の辞書があり、一番下の階層に具体的な1の地点を示す複数の辞書が存在する。

【0026】

図2乃至図12は図1に示す本発明実施形態の動作を説明するために引用した図であり、うち、図2乃至図5は、階層構造を持つ音声認識辞書の階層辞書ツリー、図6乃至図8は、動作を示すフローチャート、図9乃至図12は絞込み条件の初期設定のための画面構成を示す。

【0027】

本発明の特徴は、階層構造を持つ音声認識辞書の検索時、第1階層において第1絞り込み条件辞書（図3に示す階層辞書ツリーではジャンル名辞書）と共に、あらかじめユーザが設定した1つあるいは複数の施設名称辞書（絞り込み条件によって区分された辞書であって、図3に示す階層辞書ツリーでは病院の辞書と宿泊施設の辞書）も認識対象として認識処理を行うことにある。

【0028】

すなわち、ユーザが頻繁に使用するジャンルや地域名等絞り込み条件をあらかじめ設定しておけば、この絞り込み条件に合致する施設名称に関しては、階層構造を順に辿り絞り込み条件を決定するという煩わしい処理を行うことなく、1回の発声で目的とする施設名称を検索することができる。また、同時に絞り込み条件辞書も認識対象としているので、あらかじめ設定してある絞り込み条件に合致しない施設名称を検索する必要が生じた場合でも、従来の階層構造を順に辿り絞り込み条件を決定していく手順により検索可能である。

【0029】

あらかじめ設定しておける施設名称辞書（絞り込み条件によって区分された辞書）の数あるいはサイズは、認識率の観点からや使用可能なメモリ容量の制限などからシステム設計者があらかじめ設定しておけばよい。

【0030】

第1階層の認識処理において、ジャンル名辞書中の単語が認識された場合でも、図5に示す階層辞書ツリーのように、サブジャンル名辞書と共にユーザがあらかじめ設定した施設名称辞書（絞り込み条件によって区分された辞書であって、図5に示す階層辞書ツリーでは病院の辞書と宿泊施設の辞書）等、認識される際に認識対象の辞書の待ち受けワードとして記憶されていた待ち受けワードのうち、認識された音声に関連する待ち受けワード等、この絞り込み条件に合致する辞書（図5に示す階層辞書ツリーでは宿泊施設の辞書）も認識対象としてもよい。

第3階層以降についても同様である。

【0031】

以下、図面を参照しながら詳細に説明する。まず、図2に示した階層辞書ツリーに従えば、システムとユーザとのやりとりは以下に示すようになる。

- (1) システム「コマンドをどうぞ」
- (2) ユーザ 「病院」
- (3) システム「次のジャンルをどうぞ」
- (4) ユーザ 「診療所」
- (5) システム「都道府県名をどうぞ」
- (6) ユーザ 「埼玉県」
- (7) システム「市区町村名をどうぞ」
- (8) ユーザ 「川越市」
- (9) システム「名称をどうぞ」
- (10) ユーザ 「栗田医院」

すなわち、ここでは、入力音声「栗田医院」に対して、「埼玉県川越市の病院（診療所）の辞書」204を認識対象として音声認識を行う。

【0032】

これに対し、図3に階層辞書ツリーを示すように、本発明の特徴である、ユーザがあらかじめ「病院」302、「宿泊施設」303を設定していた場合であって、設定した絞り込み条件に合致した施設名称を検索する場合のシステムとユーザのやりとりは以下のようになる。

(1) システム「ジャンル名または施設名称をどうぞ」

(2) ユーザ「斎藤医院」

【0033】

この場合、入力音声「斎藤医院」に対して、「ジャンル名辞書」301、「病院の辞書」302、「宿泊施設の辞書」303を認識対象として音声認識が行なわれる。ここでは、目的語（斎藤医院）が、「病院の辞書」302に含まれるため、1回の発話で検索処理が終了する。「病院の辞書」302は、「全ての都道府県」の「全ての市区町村」の「病院」の全てのサブジャンルに属する名称の辞書（307、308、…313…）の集合であり、「宿泊施設の辞書」303も同様である。

【0034】

一方、図4に階層辞書ツリーを示すように、設定した絞り込み条件に合致しない施設名称を検索する場合で、第2階層以降では絞り込み条件辞書のみを認識対象とする場合のシステムとユーザのやりとりは以下のようになる。

(1) システム「ジャンル名または施設名称をどうぞ」

(2) ユーザ 「駅名」

(3) システム「サブジャンル名をどうぞ」

(4) ユーザ 「私鉄」

(5) システム「都道府県名をどうぞ」

(6) ユーザ 「埼玉県」

(7) システム「市区町村名をどうぞ」

(8) ユーザ 「熊谷市」

(9) システム「駅名をどうぞ」

(10) ユーザ 「石原駅」

【0035】

ここでは、入力音声「石原駅」に対して、「埼玉県熊谷市の駅名（私鉄）の辞書」408を認識対象として音声認識を行う。目的語（石原駅）は、第一階層待ち受け辞書400に含まれないため、ユーザは第1階層においてジャンル名辞書401に含まれるジャンル名を発話し、以後、従来手法に従い検索処理を行う。

【0036】

次に、設定した絞り込み条件に合致した施設名称を検索する場合で、第2階層以降では絞り込み条件辞書とともにあらかじめ設定した絞り込み条件、および検索の過程で決定した絞り込み条件に合致する施設名称辞書を認識対象とする場合について図5を参照しながら説明する。ここではシステムとユーザのやりとりは以下のようになる。

- (1) システム「ジャンル名または施設名称をどうぞ」
- (2) ユーザ 「宿泊施設」
- (3) システム「サブジャンル名または施設名称をどうぞ」
- (4) ユーザ 「小林旅館」

【0037】

ここでは、入力音声「小林旅館」に対して、「宿泊施設のサブジャンル名辞書」505、「宿泊施設の辞書」503を認識対象として音声認識を行う。目的語（小林旅館）が、「宿泊施設の辞書」503に含まれるので、この時点で検索処理は終了する。

【0038】

第2階層以降では絞り込み条件辞書とともにあらかじめ設定した絞り込み条件および検索の過程で決定した絞り込み条件に合致する施設名称辞書を認識対象とする。例えば、

- (1) システム「ジャンル名または施設名称をどうぞ」
- (2) ユーザ 「宿泊施設」
- (3) システム「サブジャンル名または施設名称をどうぞ」
- (4) ユーザ 「旅館」
- (5) システム「都道府県名または施設名称をどうぞ」
- (6) ユーザ 「小林旅館」

となる。

【0039】

なお、あらかじめ設定した絞り込み条件に合致しない施設名称を検索する場合のシステムとユーザのやりとりは以下のようになる。

- (1) システム「ジャンル名または施設名称をどうぞ」

- (2) ユーザ 「駅名」
- (3) システム「サブジャンル名をどうぞ」(*)
- (4) ユーザ 「JR」
- (5) システム「都道府県名をどうぞ」(*)
- (6) ユーザ 「埼玉県」
- (7) システム「市区町村名をどうぞ」(*)
- (8) ユーザ 「熊谷市」
- (9) システム「駅名をどうぞ」
- (10) ユーザ 「熊谷駅」

【0040】

ここでは、入力音声「熊谷駅」に対して、「埼玉県熊谷市の駅名（JR）の辞書」を認識対象として音声認識を行う。なお、上述したシステムとユーザのやりとりにおいて、*印が付加されたものは、あらかじめ設定した絞り込み条件と検索の過程で決定した絞り込み条件全てに合致する施設名称はないのでシステム側のガイダンスに「施設名称」は含まれない。

【0041】

図6は、図3に示す階層辞書ツリーの階層展開の手順をフローチャートで示した図である。以下、図3に示す階層辞書ツリーおよび図6に示すフローチャートを参照しながら図1に示す本発明実施形態の動作について説明する。

【0042】

まず、ステップS600でユーザが初期設定部108により絞り込み条件を設定する。初期設定値は、初期設定記憶部106に記憶されるので、この処理は初期時に一度行えば良く、検索のたびに行う必要はない。ステップS601で発話ボタン等により検索開始のトリガがかかったか否かを判断し、トリガがかかっていないければステップS601の処理へ戻る。

【0043】

一方、トリガがかかったらステップS602の処理へ進み、ジャンル名辞書301と初期設定記憶部106に記憶された、ユーザがあらかじめ設定した条件に合致する1つあるいは複数の施設名称辞書をRAM103にロードする。ステッ

STEP S603でRAM103にロードされた辞書を認識対象として認識処理を行う。この際、ユーザは、ジャンル名またはあらかじめ設定した条件に合致する施設名称を発声する。

【0044】

ステップS604で、ステップS603の認識結果が施設名称であればステップS613の処理へ進み、結果表示部110で結果表示を行い、TTS(Text To Speech)出力、検索部111による検索処理等を行う。ステップS604で認識結果が施設名称でなければステップS605の処理へ進み、認識結果のジャンルのサブジャンル名辞書をRAM103にロードする。ステップS606でユーザが発声したサブジャンル名に対してRAM103にロードされた辞書を認識対象として認識処理を行う。

【0045】

ステップS607では、都道府県名辞書をRAM103にロードし、ステップS608でユーザが発声した都道府県名に対してRAM103にロードされた辞書を認識対象として認識処理を行う。ステップS609では、ステップS608における認識結果の都道府県の市区町村名辞書をRAM103にロードし、ステップS610でユーザが発声した市区町村名に対してRAM103にロードされた辞書を認識対象として認識処理を行う。

【0046】

ステップS611では、ステップS603、ステップS606、ステップS608、ステップS610の認識結果として得られた条件に合致する施設名称辞書をRAM103にロードし、ステップS612でユーザが発声した施設名称に対してRAM103にロードされた辞書を認識対象として認識処理を行う。最後にステップS613で結果表示部110による結果表示、TTS出力、検索部111による検索処理等を行う。

【0047】

図7は、図5に示す階層辞書ツリーの階層展開の手順をフローチャートで示した図である。以下、図5に示す階層辞書ツリーおよび図7に示すフローチャートを参照しながら、図1に示す本発明実施形態の動作について説明する。

【0048】

まず、ステップS700でユーザが初期設定部108を介して絞り込み条件を設定する。初期設定値は初期設定記憶部106に記憶されるので、この処理は初期設定時に一度行えば良く、検索のたびに行う必要はない。ステップS701で発話ボタン等により検索開始のトリガがかかったか否かを判断し、トリガがかかっていないければステップS701の処理へ戻る。トリガがかかったらステップS702の処理へ進み、ジャンル名辞書と初期設定記憶部106に記憶されたユーザがあらかじめ設定した条件に合致する1つあるいは複数の施設名称辞書をRAM103にロードする。ステップS703でRAM103にロードされた辞書を認識対象として認識処理を行う。この際、ユーザはジャンル名またはあらかじめ設定した条件に合致する施設名称を発声する。

【0049】

ステップS704でステップS703の認識結果が施設名称であればステップS716の処理へ進む。認識結果が施設名称でなければステップS705の処理へ進み、認識結果のジャンルのサブジャンル名辞書とあらかじめ設定した条件及びステップS703の認識結果として得られた条件の両方に合致する施設名称辞書をRAM103にロードし、ステップS706でユーザが発声したサブジャンル名または施設名称に対してRAM103にロードされた辞書を認識対象として認識処理を行う。

【0050】

ステップS707でステップS706の認識結果が施設名称であればステップS716の処理へ進む。認識結果が施設名称でなければステップS708の処理へ進み、都道府県名辞書とあらかじめ設定した条件、及びステップS703、ステップS706の認識結果として得られた条件の全てに合致する施設名称に対してRAM103にロードし、ステップS709でユーザが発声した都道府県名または施設名称に対してRAM103にロードされた辞書を認識対象として認識処理を行う。

【0051】

ステップS710で、ステップS709の認識結果が施設名称であればステッ

PS716の処理へ進む。認識結果が施設名称でなければステップS711の処理へ進み、ステップS709の認識結果の都道府県の市区町村名辞書とあらかじめ設定した条件及びステップS703、ステップS706、ステップS709の認識結果として得られた条件の全てに合致する施設名称辞書をRAM103にロードし、ステップS712でユーザが発声した市区町村名または施設名称に対してRAM103にロードされた辞書を認識対象として認識処理を行う。

【0052】

ステップS713でステップS712の認識結果が施設名称であればステップS716の処理へ進む。認識結果が施設名称でなければステップS714の処理へ進む。認識結果が施設名称でなければステップS714の処理へ進み、ステップS703、ステップS706、ステップS709、ステップS712の認識結果として得られた条件の全てに合致する施設名称辞書をRAM103にロードし、ステップS715でユーザが発声した施設名称に対してRAM103にロードされた辞書を認識対象として認識処理を行う。最後に、ステップS716で結果表示、TTS出力、検索処理等を行う。

【0053】

図8は、図6及び図7に示した認識処理（ステップS603、ステップS606、ステップSS608、ステップS610、ステップS612、ステップS703、ステップS706、ステップS709、ステップS712、ステップS715）の詳細手順をフローチャートで示した図である。

【0054】

以下、図8に示すフローチャートを参照しながら上記各ステップで実行される認識処理について説明する。まず、ステップS800で、マイク100からの入力に音声が含まれているかを検出する。検出方法として、パワーがある閾値を越えた場合に音声とみなす等の方法がある。音声が検出されれば音声開始と判断し、ステップS801で特徴量計算部101による特徴量の計算が行なわれ、ステップS802でRAM103にロードされた認識辞書に含まれる各単語と入力音声から計算された特徴量との類似度を計算する。ステップS803で音声が終了していなければステップS801の処理へ戻る。音声が終了していたら、ステッ

PS804で類似度が一番高い単語を認識結果として出力する。

【0055】

最後に、絞込み条件の初期設定の方法について、リモコンを用いて行う場合と、音声認識により行う2例を説明する。

【0056】

リモコンで行う場合、まず、リモコンのメニューボタンを押すことによって表示されるメニュー画面から「絞り込み条件設定変更」の項目を選択する。このことにより、図9に示す絞り込み条件設定変更画面となる。絞り込み条件設定変更画面は、絞込み条件（都道府県名、ジャンル名）によって区分された施設名辞書群がマトリクス状に割り付けられ配置されている。ここでは、リモコンのジョイスティックによりカーソルを移動し、設定変更を行いたい条件名に移動する。

【0057】

例えば、ジョイスティックを横方向に動かすことで、図10に示すように、都道府県一覧の中の所望の県を選択する。ここで、例えば埼玉県を選択している時にリモコンの決定ボタンを押すことによりカーソル位置の条件（埼玉県に存在する全てのジャンルの施設名称辞書）を絞り込み条件とする。

【0058】

また、例えばジョイスティックを縦方向に動かすことにより、図11に示すようにジャンル名一覧中の所望のジャンルを選択する。ここで、例えば病院を選択している時に決定ボタンを押すことによりカーソル位置の条件（全国の病院の名称辞書）を絞り込み条件とする。更に、図10に示す表示画面を用いて埼玉県を選択した後に、図11に示すように病院を選択すると、図12に示すように埼玉県にある病院の名称辞書を絞り込み条件とする。

【0059】

なお、ここでは、初期設定値として、「埼玉県」と「病院」が設定された場合に選択される名称辞書が示されているが、都道府県名と病院名が共に設定されることとは必須でなく、それぞれ単独に設定されてもよい。また、予め決定ボタンを押された位置の条件が絞り込み条件となっている場合は、絞り込み条件であることを解除するものとする。即ち、予め絞込条件となっている場合はその条件を解

除し、条件となっていない場合は絞込条件に設定を変更する。更に、上記はジョイティックで絞込みの条件を選択する例を示したが、ジョイティックの代わりにタッチパネルで行ってもよい。

【0060】

絞込み条件の初期設定を音声認識で行う例について説明する。音声認識の第1階層待ち受け辞書に「設定変更」等の絞り込み条件変更処理を意味する単語も含めておき、この単語が認識された場合には絞り込み条件設定変更処理を開始する。まず、設定変更処理において、絞り込み条件名を待ち受け単語とした辞書を用いて音声認識処理を行い、認識された条件がオンの場合はその条件をオフに、オフの場合はオンに設定を変更する。

【0061】

次に、設定変更処理において、各絞り込み条件名の後に「をオンにする」「をオフにする」等を付加したものを持ち受け単語とした辞書を用いて音声認識処理を行い、認識された単語が「(条件名)をオンにする」の場合はその条件をオンに、「(条件名)をオフにする」の場合はオフに設定する。なお、先の設定変更処理において、「(条件名)」+「(オン、オフを指定する単語)」という文法を用いた連続認識としてもよい。

【0062】

以上説明のように本発明は、最上位階層において絞り込み条件の辞書の他に、あらかじめユーザにより設定される少なくとも一つの絞込み条件によって区分された辞書を認識の対象として認識処理を行うことにより操作性の改善がはかれると共に、応答性の向上もはかれるものである。

【0063】

【発明の効果】

以上説明のように本発明によれば、階層構造を持つ音声認識辞書を複数持つ音声認識装置に用いられ、最上位階層において絞り込み条件の辞書と共に、あらかじめユーザにより設定される少なくとも一つの絞込み条件によって区分された辞書を認識の対象として認識処理を行うことにより操作性の改善ならびにレスポンスの向上がはかれ、また、ユーザが頻繁に使用する絞り込み条件(ジャンルや地

域名等)をあらかじめ設定しておくことにより、この絞り込み条件に合致する施設名称に関しては、階層構造を順に辿り絞り込み条件を決定するという煩わしい処理を行うことなく、1回の発声で目的とする施設名称を検索することができる。

【0064】

また、本発明によれば、あらかじめ設定された絞り込み条件に合致しない施設名称を検索する場合には、絞り込み条件を順に決定していくという従来の手順を踏むことができる。更に、あらかじめ設定された絞り込み条件に合致する施設名称を検索する場合でも、従来の手順で絞り込み条件を順に決定した上で最終的に絞り込み条件に合致する1つの辞書セットを用いて施設名称の認識処理を行うことも可能である。

【図面の簡単な説明】

【図1】

本発明における音声認識装置の一実施形態を示すブロック図である。

【図2】

本発明の音声認識方法を説明するために引用した図であり、階層辞書ツリーの一例を示す図である。

【図3】

本発明の音声認識方法を説明するために引用した図であり、階層辞書ツリーの一例を示す図である。

【図4】

本発明の音声認識方法を説明するために引用した図であり、階層辞書ツリーの一例を示す図である。

【図5】

本発明の音声認識方法を説明するために引用した図であり、階層辞書ツリーの一例を示す図である。

【図6】

図3に示す階層辞書ツリーの階層展開の手順をフローチャートで示した図である。

【図7】

図5に示す階層辞書ツリーの階層展開の手順をフローチャートで示した図である。

【図8】

図6、図7における認識処理の処理手順の詳細をフローチャートで示した図である。

【図9】

絞込み条件の初期設定方法を表示画面上に示した図である。

【図10】

絞込み条件の初期設定方法を表示画面上に示した図である。

【図11】

絞込み条件の初期設定方法を表示画面上に示した図である。

【図12】

絞込み条件の初期設定方法を表示画面上に示した図である。

【図13】

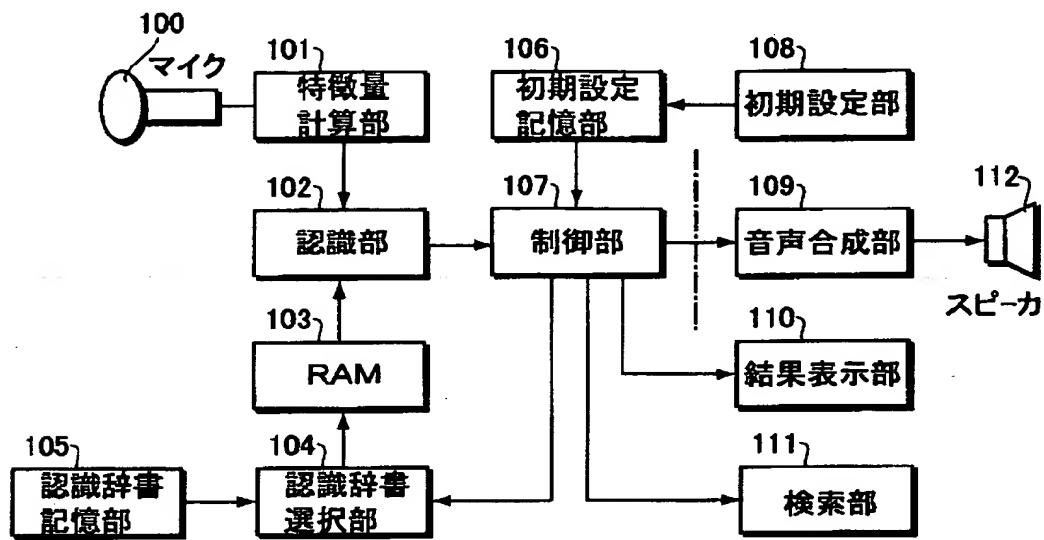
従来における絞込みの手順を説明するために引用した図である。

【符号の説明】

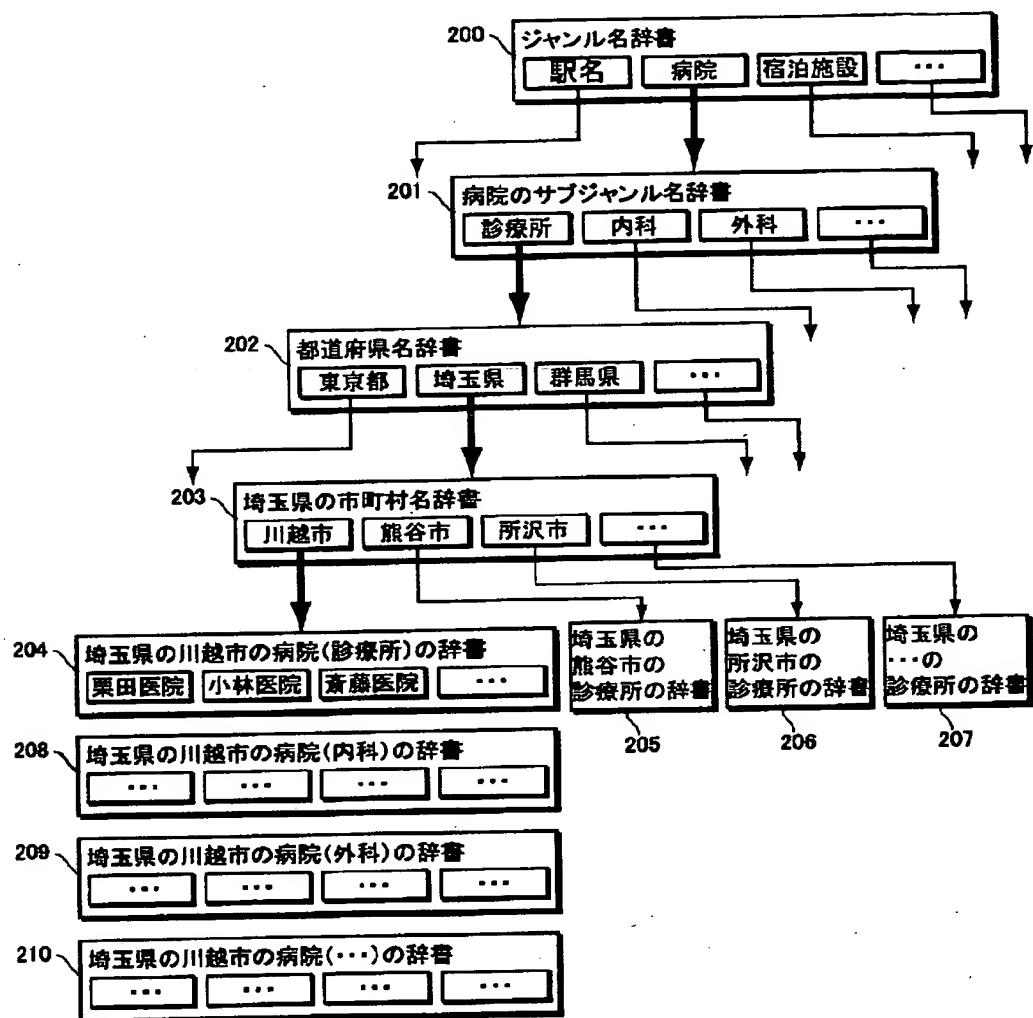
100…マイク、101…特微量計算部、102…認識部、103…RAM、104…認識辞書選択部、105…認識辞書記憶部、106…初期設定記憶部、107…制御部、108…初期設定部、109…音声合成部、110…結果表示部、111…検索部、112…スピーカ

【書類名】 図面

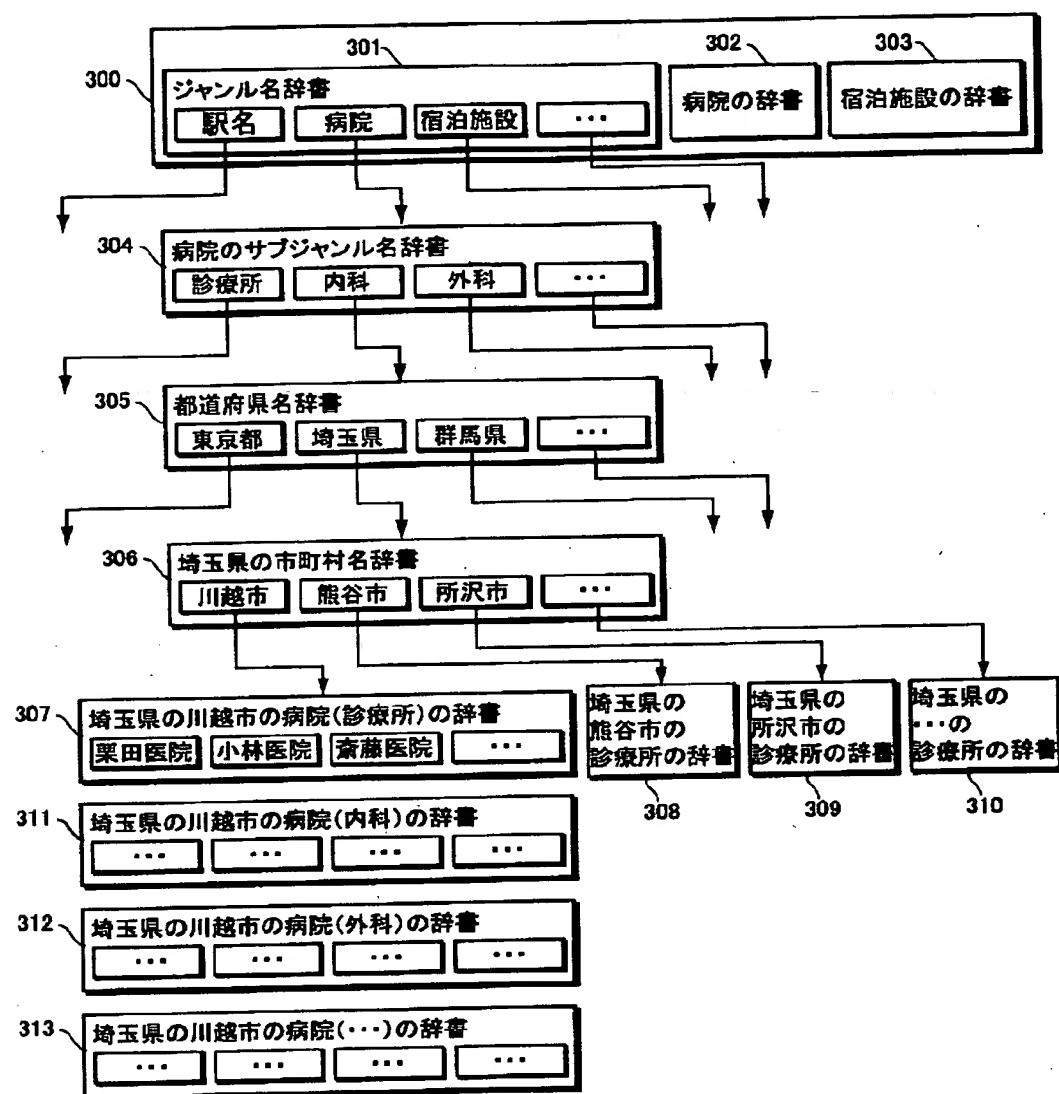
【図1】



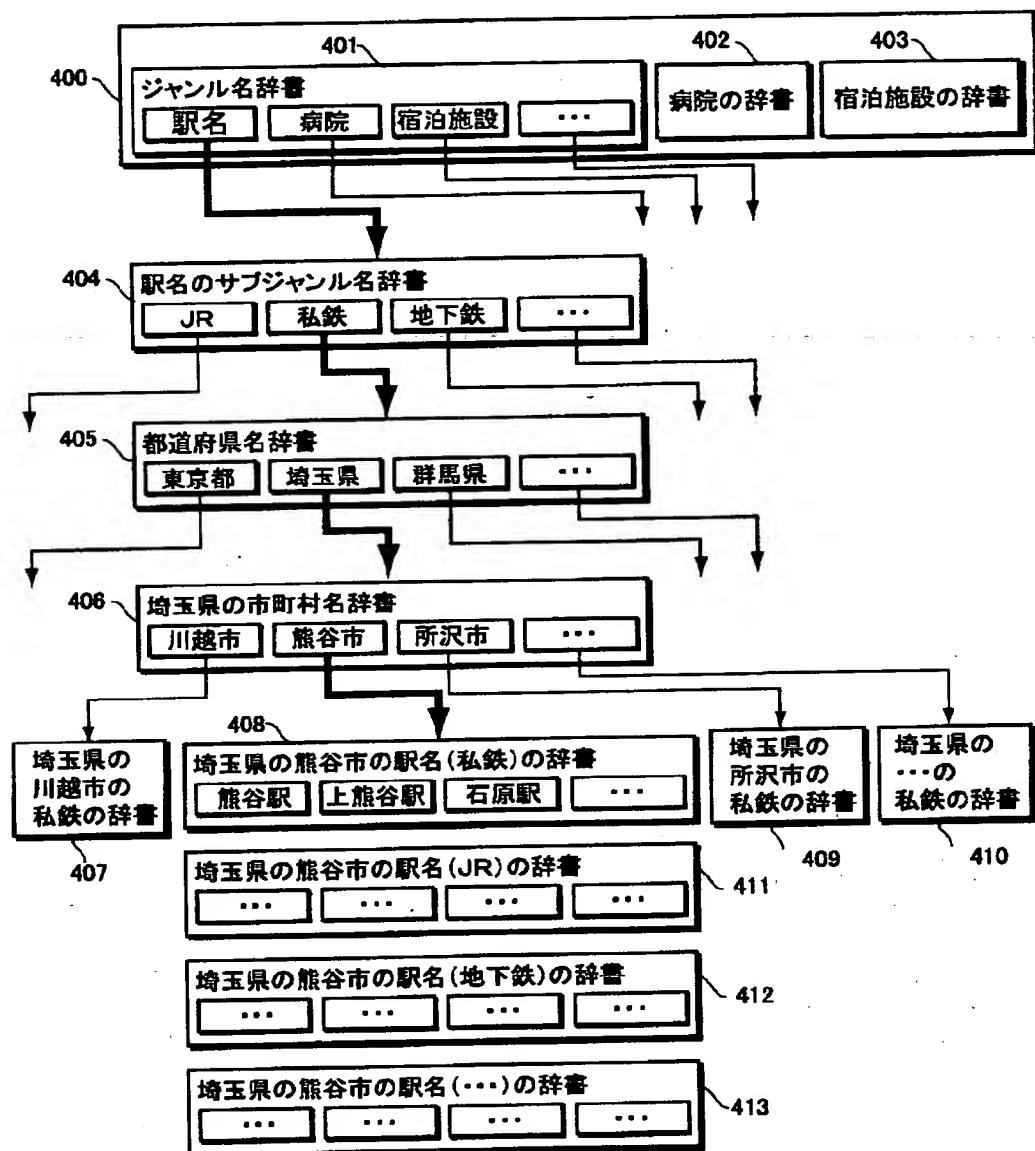
【図2】



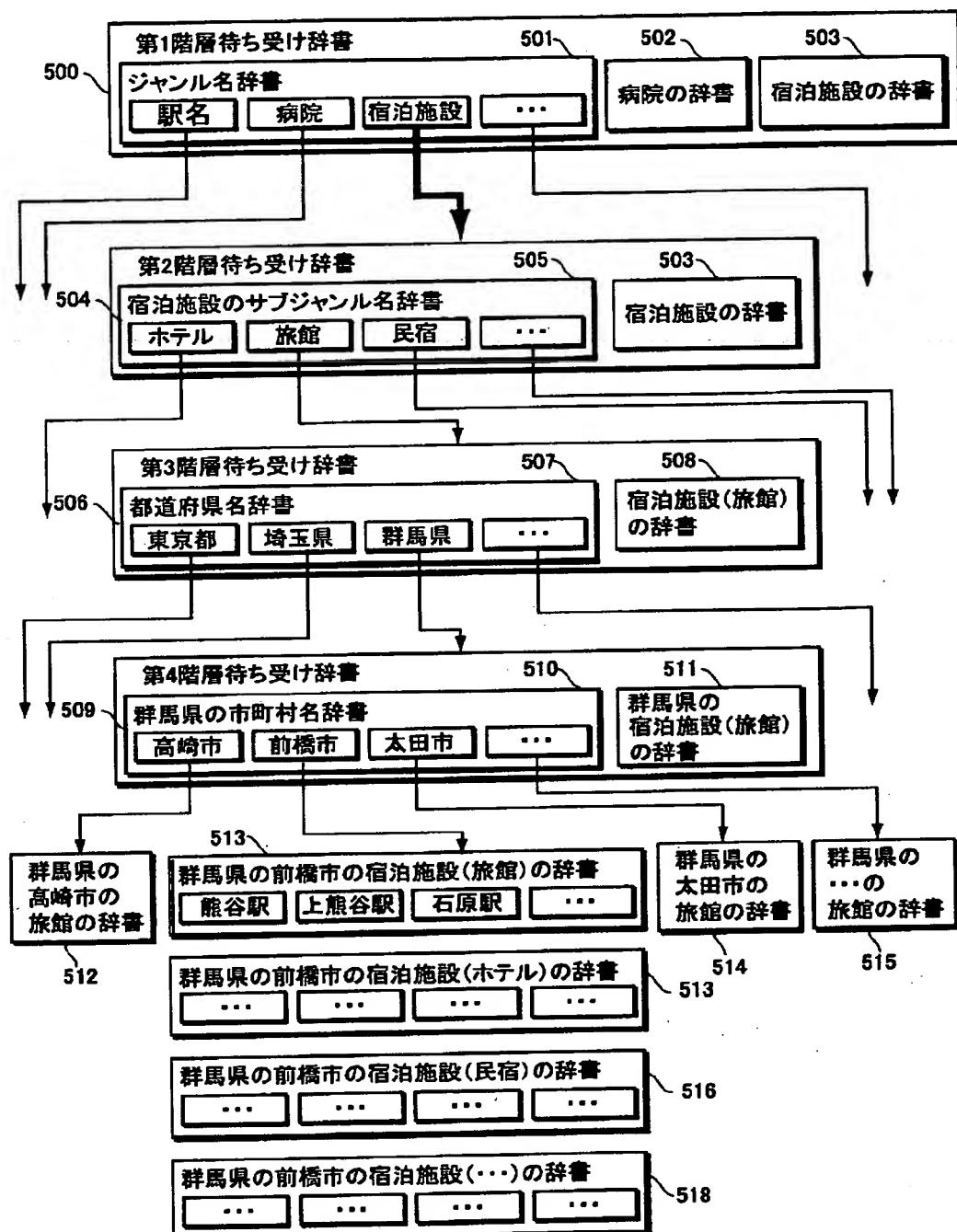
〔四三〕



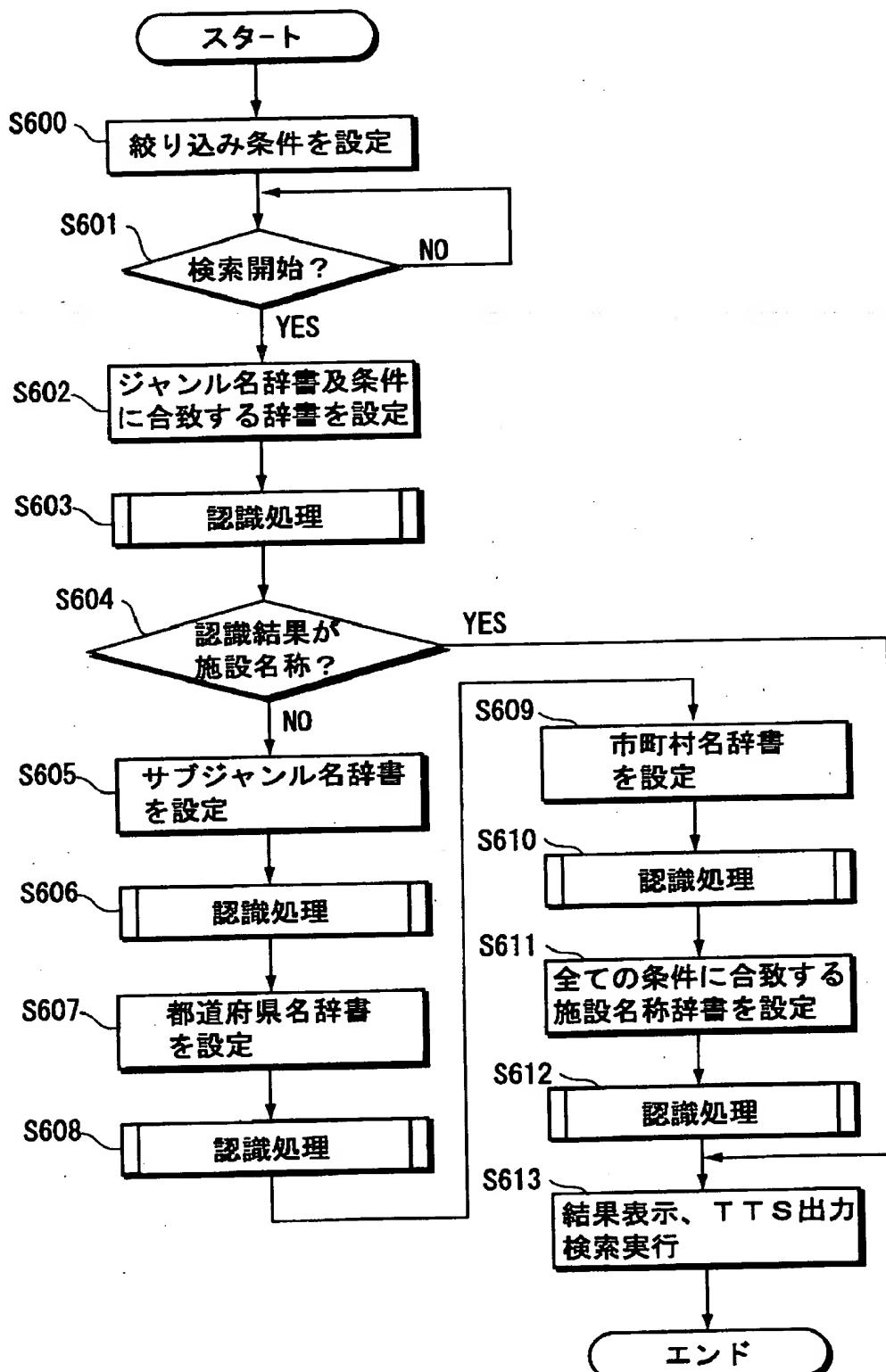
【図4】



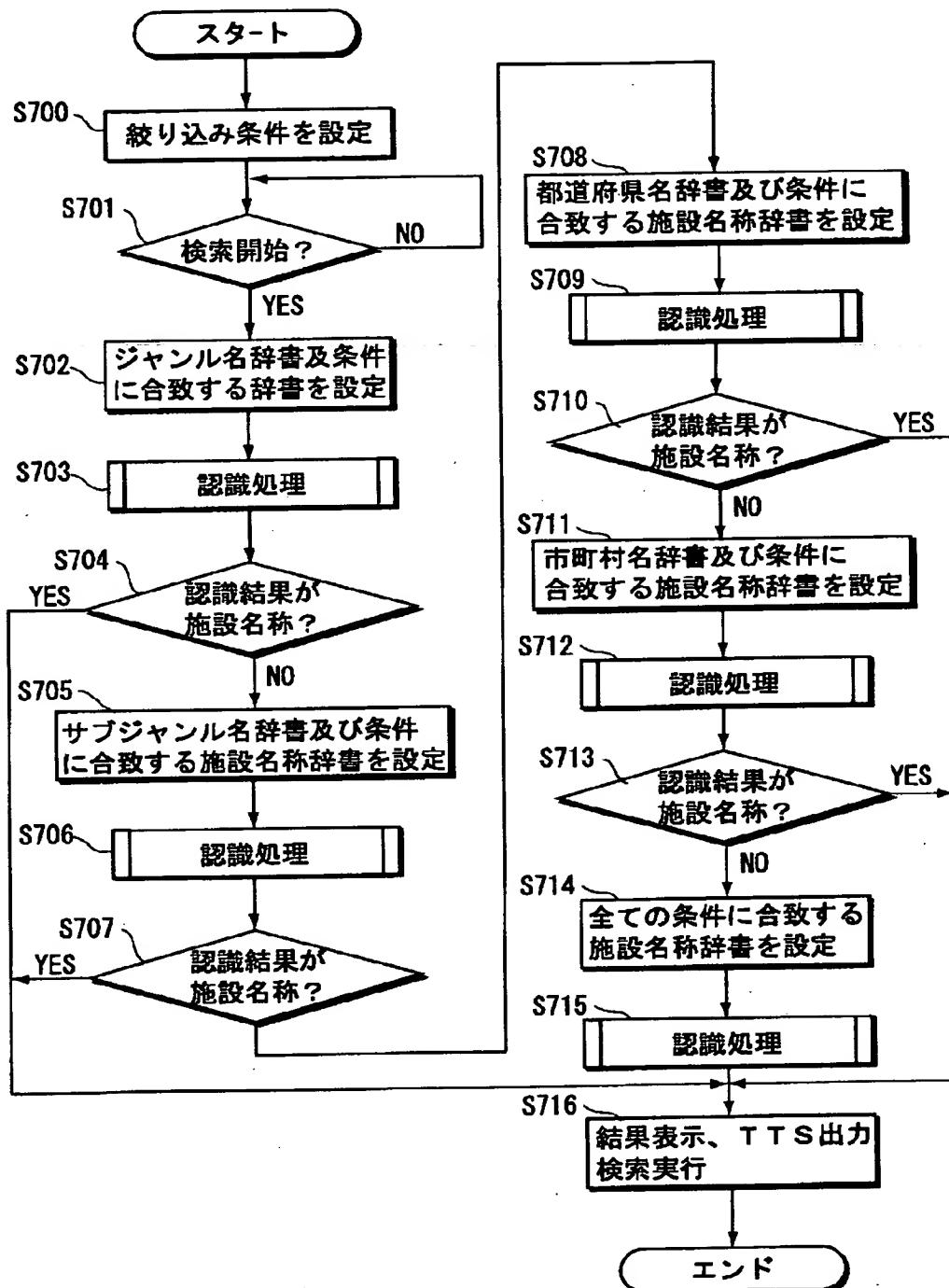
【図5】



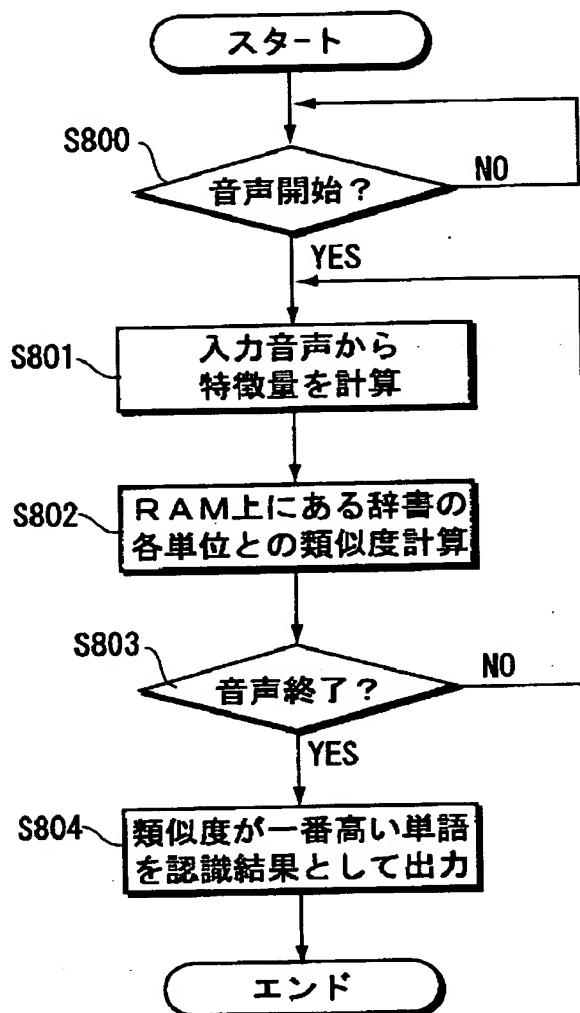
【図6】



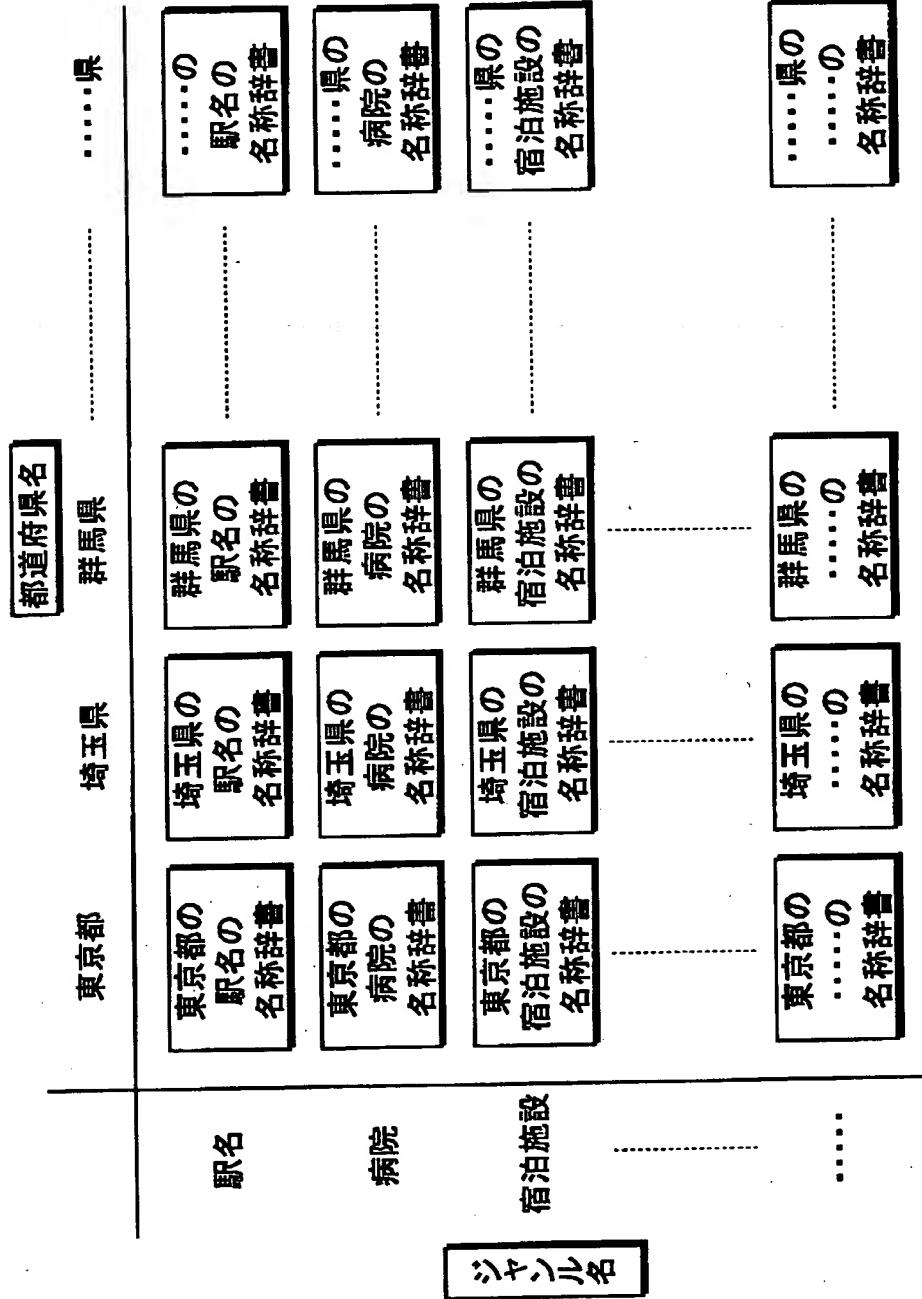
【図7】



【図8】

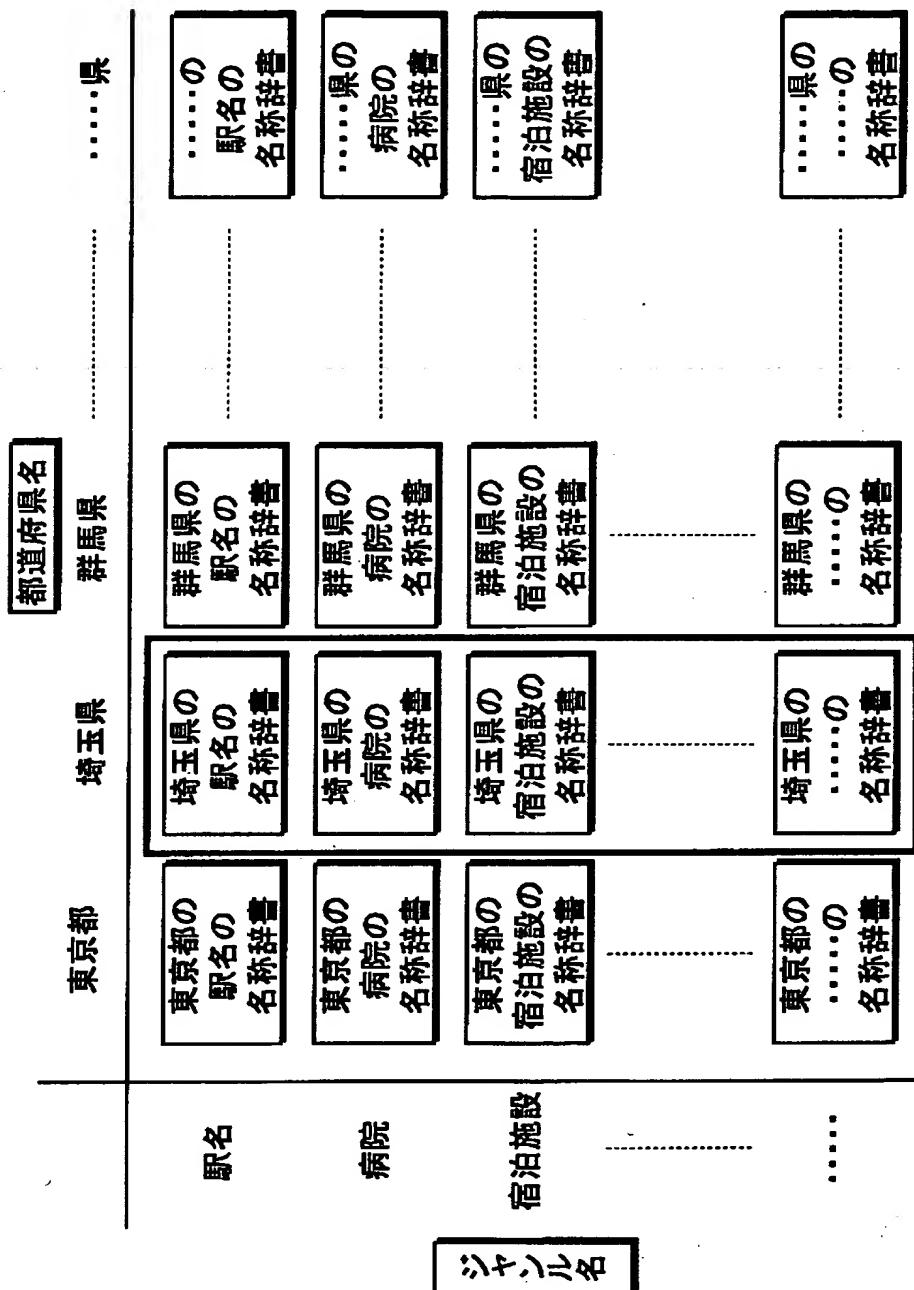


【図9】



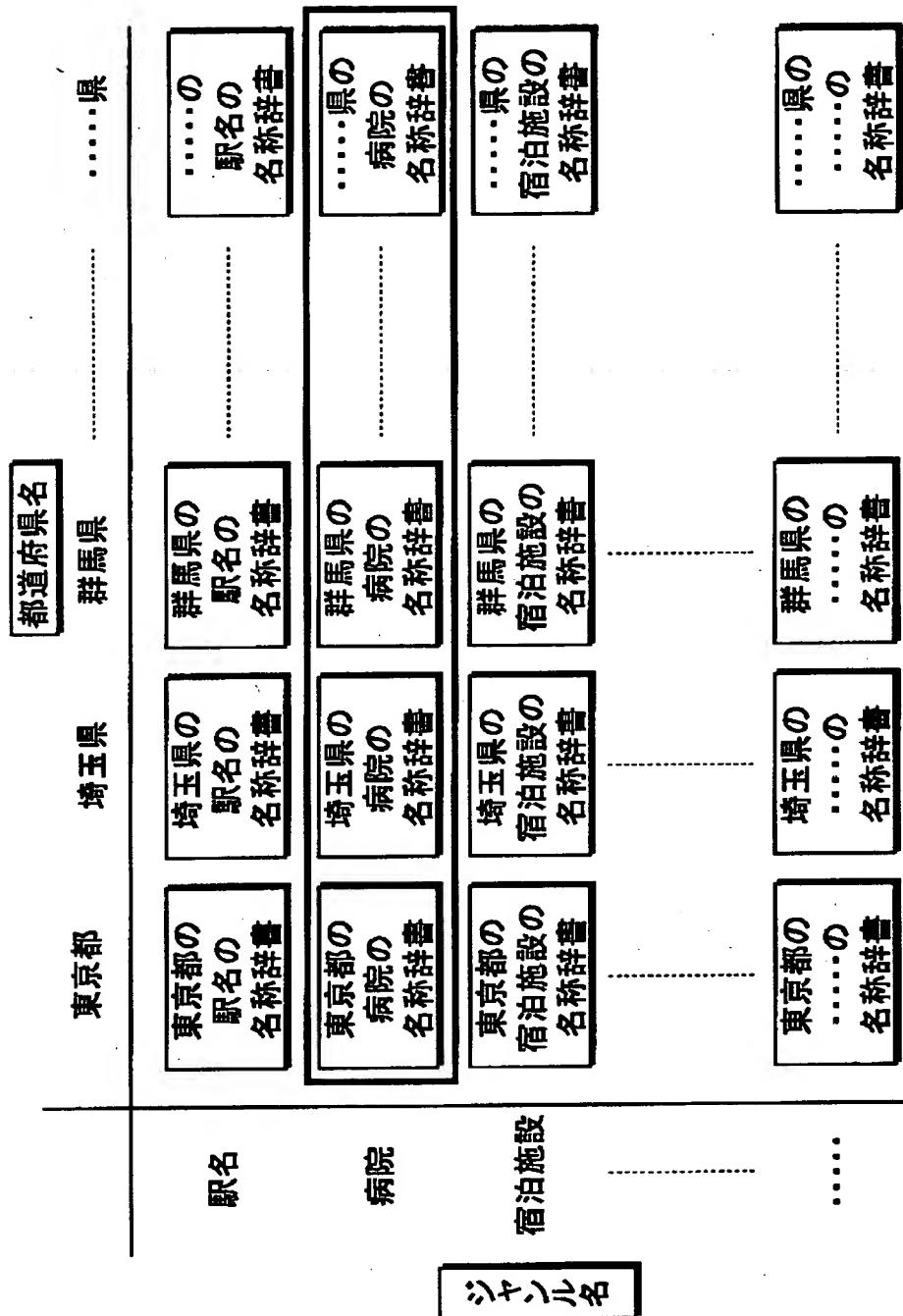
取り込み条件(都道府県名、ジャンル名)によって区分された施設名称辞書群

【図10】



初期設定として「埼玉県」が選択された場合に選択される名前辞書都

【図11】



初期設定として「病院」が設定された場合に選択される名前辞書

【図12】

初期設定として「埼玉県」と「病院」が設定された場合に選択される名称辞書

【図13】

- ① システム：「ジャンルの名称をどうぞ」
- ② ユーザ：「教育施設」
- ③ システム：「次のジャンルの名称をどうぞ」
- ④ ユーザ：「高校」
- ⑤ システム：「都道府県名をどうぞ」
- ⑥ ユーザ：「東京都」
- ⑦ システム：「市区町村名をどうぞ」
- ⑧ ユーザ：「渋谷区」
- ⑨ システム：「名称をどうぞ」
- システム：「〇〇学園」

【書類名】 要約書

【要約】

【課題】 施設名称等を検索するときに、階層構造を持つ音声認識辞書の階層構造を順に辿って全ての絞り込み条件を決定するという煩わしい手順の実行を回避し、操作性と応答性の向上をはかる。

【解決手段】 本発明の音声認識装置は、階層構造を持つ複数の音声認識辞書（認識辞書記憶部105）と、音声認識辞書のうち所望の辞書を待ち受けワード一覧として抽出する抽出手段（制御部107）と、音声認識辞書のうち、所望の辞書の選択操作を行う選択手段（認識辞書選択部104）と、前記選択手段で選択された辞書を最上位階層の待ち受けワード一覧とし、前記抽出手段で抽出された通常の辞書と併せて記憶する記憶手段（RAM103）と、入力された音声と記憶手段に記憶された待ち受けワード一覧とを比較することによって入力された音声を認識する認識手段（認識部102）で構成される。

【選択図】 図1

出願人履歴情報

識別番号 [000005016]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都目黒区目黒1丁目4番1号

氏 名 パイオニア株式会社